Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Journal of Multivariate Analysis 95 (2005) 410–430
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

Noncentral quadratic forms of the skew elliptical variables

B.Q. Fang[1]

Institute of Applied Mathematics, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing 100080, China

Received 8 April 2003
Available online 29 September 2004

Abstract

In this paper the quadratic forms in the skew elliptical variables are studied. A family of the noncentral generalized Dirichlet distributions is introduced and their distribution functions and probability density functions are obtained. The moment generating functions of the quadratic forms in the skew normal variables are obtained. Sufficient and necessary conditions for the quadratic forms in the skew normal variables to have the noncentral generalized Dirichlet distributions are obtained. This leads to the noncentral Cochran's Theorem for the skew normal distribution.

AMS 2000 subject classification: primary 62H10; secondary 62E15

Keywords: Normal distribution; Skew normal distribution; Skew elliptical distribution; Quadratic forms; Cochran's Theorem

1. Introduction

Quadratic forms in random variables arise in many problems of statistical analysis. Many statistics can be expressed as functions of quadratic forms, including the χ^2 statistic, the two-sided t-statistic and the F statistic as examples. Let $\mathbf{x}\sim N(\mu, \Sigma)$ and $A_i = A_i'$, $i = 1, 2$, symmetric. Two main properties of the quadratic forms in normal variables are the necessary and sufficient conditions for $\mathbf{x}' A_1 \mathbf{x}$ to have noncentral chi-square distribution

E-mail address: fangbiqi@amss.ac.cn.

[1] Supported by the National Science Foundation of China.

0047-259X/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2004.08.009

and the necessary and sufficient conditions for $\mathbf{x}'A_i\mathbf{x}$, $i = 1, 2$, to be independent, see e.g. Rao [19].

The skew normal distributions and the skew elliptical distributions are extensions of the normal distributions and the elliptical distributions. These distributions provide a more flexible way to model data presenting skewness and still have a kind of symmetry similar to the normal and elliptical distributions to maintain tractability in analysis. They have received increasing attention by many authors recently. The distributions of the quadratic forms of the skew normal variables in the central case (with location zero) were studied, for example, in Azzalini and Dalla Valle [5], Azzalini and Capitanio [6], Loperfido [16]. A version of Cochran's Theorem for the skew normal distribution in the central case was obtained in Azzalini and Capitanio [6] and a version of Cochran's Theorem for the skew elliptical distribution in the central case was obtained in Fang [11].

In this paper we investigate the quadratic forms in the skew elliptical variables in the noncentral case with emphasis on the quadratic forms in the skew normal variables. We obtain two versions of Cochran's Theorem for the noncentral quadratic forms in the skew normal variables. In Section 2, the distribution functions and the probability density functions of the noncentral quadratic forms in the skew elliptical variables are obtained. The definition of the noncentral generalized Dirichlet distributions is given. These distributions include the chi-square distribution, the so-called G distribution (central generalized Dirichlet distribution) in Anderson and Fang [2] and generalized χ^2 distribution in Fan [10] as special cases. In Section 3, the moment generating functions and the first two moments of the noncentral quadratic forms in the skew normal variables are obtained. Section 4 gives the necessary and sufficient conditions for the joint distribution of several quadratic forms in the skew normal variables to be the noncentral generalized Dirichlet distribution. This leads to two versions of noncentral Cochran's Theorem for the skew normal distributions. Section 5 provides two examples. In the first example, an estimate for the measure of multivariate kurtosis defined by Mardia [17] is obtained for a real data set. The second example is the linear model where the error vector has the skew elliptical distribution. We investigate the effect of the skew parameters on the least-square estimator of the location parameter. Numerical results of the true level of the conventional confidence interval for the scale parameter based on normality assumption are obtained. The proofs of the propositions and theorems are collected in Section 6. A brief discussion is given in Section 7. For convenience of the reader, we give the definition of the skew elliptical distribution and some of its properties in Fang [11] as follows.

Definition 1. Let f be the density generator of an n-dimensional spherical distribution, satisfying $\int_{R^n} f(\mathbf{v}'\mathbf{v})\,d\mathbf{v} = 1$, F_1 its one-dimensional marginal distribution function, $\lambda \in R$, $\alpha \in R^k$ and $\xi \in R^k$ be constant and Ω a $k \times k$ constant positive definite matrix, $k = n-1$. Let $\mathbf{z} \in R^k$ be a random vector with probability density function

$$\int_{-\infty}^{\lambda+\alpha'(\mathbf{z}-\xi)} f(y_0^2 + (\mathbf{z}-\xi)'\Omega^{-1}(\mathbf{z}-\xi))\,dy_0 |\Omega|^{-1/2}/F_1(\lambda/c_0), \quad \mathbf{z} \in R^k, \tag{1}$$

where $c_0 = (1 + \alpha'\Omega\alpha)^{1/2}$. Then $\mathbf{z}$ is called to have the skew elliptical distribution and denoted by $\mathbf{z} \sim S_k(\xi, \Omega, \lambda, \alpha; f)$, see Fang [11].

The family of the skew elliptical distributions is closed under linear transformation, marginalization and conditioning. Specifically, if A is a nonsingular $k \times k$ matrix, then $A'\mathbf{z} \sim S_k(A'\xi, A'\Omega A, \lambda, A^{-1}\alpha; f)$. As a result of this property, we assume $\Omega = I$ in some places of this paper for notation simplicity and the results obtained can be extended for the general $\Omega > 0$ easily. Partition $\mathbf{z}$, ξ, α and Ω as

$$\mathbf{z} = \begin{pmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{pmatrix}, \quad \xi = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}, \quad \alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}, \quad \Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}, \tag{2}$$

where $\mathbf{z}_i$, ξ_i and α_i are $k_i \times 1$ and Ω_{11} is $k_1 \times k_1$. Then

$$\mathbf{z}_1 \sim S_{k_1}(\xi_1, \Omega_{11}, c_1\lambda, c_1(\alpha_1 + \Omega_{11}^{-1}\Omega_{12}\alpha_2); f_{k_1+1}), \tag{3}$$

where $f_{k_1+1}(v_0^2 + \mathbf{v}_1'\mathbf{v}_1)$ is the k_1+1-dimensional marginal density of $f(v_0^2 + \mathbf{v}'\mathbf{v})$, $\Omega_{22.1} = \Omega_{22} - \Omega_{21}\Omega_{11}^{-1}\Omega_{12}$ and $c_1 = (1 + \alpha_2'\Omega_{22.1}\alpha_2)^{-1/2}$.

Denote by $\phi_n(x)$ the normal density generator, i.e., $\phi_n(x) = \exp(-x/2)(2\pi)^{-n/2}$. Some special cases for the skew elliptical distributions are the elliptical distribution with $\alpha = 0$, the skew normal distribution with $f = \phi_n$ in Azzalini and Dalla Valle [5], Azzalini and Capitanio [6], Arnold and Beaver [4], the skew elliptical distribution with $\lambda = 0$ in Branco and Dey [7].

2. Distribution and density functions

In this section we obtain the distribution functions and the probability density functions of the quadratic forms in the skew elliptical variables. This leads naturally to the introduction of a new multivariate distribution, the generalized Dirichlet distribution. Basic properties are investigated.

Proposition 1. *Assume* $\mathbf{z} \sim S_k(\xi, I, \lambda, \alpha; f_n)$ *and* $Q = \mathbf{z}'\mathbf{z}$. *Let* $\tau_{11} = \xi'\xi$, $\tau_{12} = \xi'\alpha$, $\tau_{22} = \alpha'\alpha$. *Let* $l_1 = \tau_{12}\tau_{22}^{-\frac{1}{2}}$, *if* $\tau_{22} \neq 0$ *or* $\tau_{11}^{\frac{1}{2}}$, *otherwise, for* $k = 1$; $l_1 = \tau_{12}\tau_{22}^{-\frac{1}{2}}$, *if* $\tau_{22} \neq 0$ *or* 0, *otherwise*, $l_2 = (\tau_{11} - l_1^2)^{\frac{1}{2}}$, *for* $k \geqslant 2$; *and* $c(k) = \pi^{\frac{k-2}{2}}/\Gamma(\frac{k-2}{2})$ *for* $k \geqslant 3$. *Then the probability density function of Q is, if* $k = 1$,

$$f_Q(x) = \frac{x^{-\frac{1}{2}}}{2F_1(\frac{\lambda}{c_0})} \left\{ \int_{-\infty}^{\lambda + \tau_{22}^{\frac{1}{2}} x^{\frac{1}{2}} - \tau_{12}} f_n(y_0^2 + x - 2l_1 x^{\frac{1}{2}} + \tau_{11})\, dy_0 \right.$$
$$\left. + \int_{-\infty}^{\lambda - \tau_{22}^{\frac{1}{2}} x^{\frac{1}{2}} - \tau_{12}} f_n(y_0^2 + x + 2l_1 x^{\frac{1}{2}} + \tau_{11})\, dy_0 \right\}; \tag{4}$$

if $k=2$,

$$f_Q(x) = \frac{1}{2F_1(\frac{\lambda}{c_0})} \int_0^{2\pi} \left\{ \int_{-\infty}^{\lambda+\tau_{22}^{\frac{1}{2}} x^{\frac{1}{2}} \cos\theta - \tau_{12}} \times f_n(y_0^2 + x - 2x^{\frac{1}{2}}(l_2 \sin\theta + l_1 \cos\theta) + \tau_{11})\, dy_0 \right\} d\theta; \tag{5}$$

if $k \geqslant 3$,

$$f_Q(x) = \frac{c(k) x^{\frac{k}{2}-1}}{2F_1(\frac{\lambda}{c_0})} \int_{\substack{0<\theta_1<\pi \\ 0<\theta_2<2\pi}} (\sin\theta_1)^{k-2} |\sin\theta_2|^{k-3} \left\{ \int_{-\infty}^{\lambda+\tau_{22}^{\frac{1}{2}} x^{\frac{1}{2}} \cos\theta_1 - \tau_{12}} \times f_n(y_0^2 + x - 2x^{\frac{1}{2}}(l_2 \sin\theta_1 \cos\theta_2 + l_1 \cos\theta_1) + \tau_{11})\, dy_0 \right\} d\theta_1\, d\theta_2, \tag{6}$$

where c_0 is *given in Definition* 1 *with* $\Omega = I$.

Proposition 2. *Assume* $\mathbf{z} \sim S_k(\xi, I, \lambda, \alpha;\, f_n)$. *Partition* $\mathbf{z}$ *into h parts as* $\mathbf{z} = (\mathbf{z}_1', \ldots, \mathbf{z}_h')'$, *where* $\mathbf{z}_i$ *is* k_i*-dimensional. Partition* ξ *and* α *in the same manner. Let* $Q_i = \mathbf{z}_i'\mathbf{z}_i$, $Q = (Q_1, \ldots, Q_h)$, $\tau_{i,11} = \xi_i'\xi_i$, $\tau_{i,12} = \xi_i'\alpha_i$, $\tau_{i,22} = \alpha_i'\alpha_i$. *Let* l_{i1}, l_{i2} *be defined as in Proposition* 1 *for* Q_i *according to* $k_i = 1, 2$ *or* $k_i \geqslant 3$. *Let* $\mathbf{y}_i$ *be of dimension* $\min(k_i, 3)$ *with components* y_{ij}. *Let* $c(k_i) = 1$, $b(\mathbf{y}_i, k_i) = I_{(0,x_i)}(\mathbf{y}_i'\mathbf{y}_i)$ *for* $k_i = 1$ *or* 2 *and* $b(\mathbf{y}_i, k_i) = I_{(0,x_i)}(\mathbf{y}_i'\mathbf{y}_i)|y_{i3}|^{k_i-3}$ *and* $c(k_i)$ *as in Proposition* 1 *for* $k_i \geqslant 3$. *Denote* $k_i' = \min(k_i, 2)$. *Then the distribution of Q is*

$$\begin{aligned} &\mathrm{pr}(Q_1 < x_1, \ldots, Q_h < x_h) \\ &\quad = \prod_{i=1}^h c(k_i) \int \prod_{i=1}^h b(\mathbf{y}_i, k_i) \left[\int I_{(-\infty,\ \lambda+\sum_{i=1}^h (\tau_{i,22}^{1/2} y_{i1} - \tau_{i,12}))}(y_0) \right. \\ &\qquad \left. \times f_n\left(y_0^2 + \sum_{i=1}^h \mathbf{y}_i'\mathbf{y}_i - 2\sum_{i=1}^h \sum_{j=1}^{k_i'} l_{ij} y_{ij} + \sum_{i=1}^h \tau_{i,11} \right) dy_0 \right] \\ &\qquad \times \prod_{i=1}^h d\mathbf{y}_i / F_1(\lambda/c_0), \end{aligned} \tag{7}$$

where c_0 *is given in Definition* 1 *with* $\Omega = I$.

The probability density function of Q can be obtained from Proposition 2, we omit these complicated formulas. It can be seen that the distribution of Q depends on ξ and α only through $\xi_i'\xi_i$, $\alpha_i'\xi_i$ and $\alpha_i'\alpha_i$.

Definition 2. The distribution of $Q = (Q_1, \ldots, Q_h)$ in (7) is called the noncentral generalized Dirichlet distribution with parameters $k_1/2, \ldots, k_h/2$, $\tau_{i,11}$, $\tau_{i,12}$, $\tau_{i,22}$, $i = 1, \ldots, h$, λ and density function generator f_n, where k_i are positive integers, $\tau_{i,11} \geqslant 0$, $\tau_{i,22} \geqslant 0$, $\sum_{i=1}^{h} k_i = k$, $n = k + 1$, $\tau_{i,12}^2 \leqslant \tau_{i,11}\tau_{i,22}$, for $k_i \geqslant 2$ and $\tau_{i,12}^2 = \tau_{i,11}\tau_{i,22}$ for $k_i = 1$, $i = 1, \ldots, h$, and denoted by $NG_h(k_1/2, \ldots, k_h/2; \tau, \lambda; f_n)$.

In some special cases, if the density generator f is specified, more compact form of the density function can be obtained. Suppose now $f = \phi_n$. Denote by Φ the distribution function of the standard normal variable and f_0 the density of the usual χ_k^2 distribution. Integrating y_0 in Proposition 1 and applying Lemma 1.3.2 in Muirhead [18, p. 21] for the case $k \geqslant 3$, we obtain the following expressions for the density function of the quadratic form in the skew normal vector.

$$f_Q(x) = f_0(x) \cdot \frac{\exp(-\tau_{11}/2)}{2\Phi(\lambda/c_0)} \{\Phi(\lambda + \tau_{22}^{\frac{1}{2}} x^{\frac{1}{2}} - \tau_{12}) \exp(l_1 x^{1/2}) + \Phi(\lambda - \tau_{22}^{\frac{1}{2}} x^{\frac{1}{2}} - \tau_{12}) \exp(-l_1 x^{1/2})\}, \quad (k = 1);$$

$$f_Q(x) = f_0(x) \cdot \frac{\exp(-\tau_{11}/2)}{2\pi\Phi(\lambda/c_0)} \int_0^{2\pi} \Phi(\lambda + \tau_{22}^{\frac{1}{2}} x^{\frac{1}{2}} \cos\theta - \tau_{12}) \times \exp(x^{\frac{1}{2}}(l_2 \sin\theta + l_1 \cos\theta))\, d\theta, \quad (k = 2);$$

$$f_Q(x) = f_0(x) \cdot \frac{\Gamma(k/2)\exp(-\tau_{11}/2)}{\pi^{1/2}\Gamma((k-1)/2)\Phi(\lambda/c_0)} \int_0^{\pi} (\sin\theta_1)^{k-2} \exp(x^{1/2} l_1 \cos\theta_1) \times \Phi(\lambda + \tau_{22}^{\frac{1}{2}} x^{\frac{1}{2}} \cos\theta_1 - \tau_{12}) {}_0F_1\left(\frac{k-1}{2}; \frac{x l_2^2 (\sin\theta_1)^2}{4}\right) d\theta_1, \quad (k \geqslant 3),$$

where ${}_pF_q$ is the generalized hypergeometric function, see Muirhead [18, p. 20]. These formulas generalize the probability density function of the noncentral chi-square distribution in Muirhead [18, Theorem 1.3.4].

Proposition 2 presents the distribution function in its most general form and is the basis for the introduction of the new distribution in Definition 2. The formula of the distribution function can be simplified in some important special cases. If $h = 1$ and $f_n = \phi_n$, then the dimension of integration space is three at most. Theorems 1 and 2 in Section 4 provide conditions that the joint distribution function can be factorized for $h \geqslant 2$. If some of the skew parameters vanish, then by the method in the proof of Proposition 1, the dimension can be reduced further. This will be illustrated in Section 5. In many applications, we use only one or two quadratic forms to form statistics and the quadratic forms are centralized. Hence the distribution function is simpler than the general formula in Proposition 2 and numerical calculation is feasible.

The family of the noncentral generalized Dirichlet distributions includes some distributions studied in the literature. If $\lambda = 0$ and $\xi = 0$, then the distribution of Q does not depend on α and can be obtained by letting $\alpha = 0$, see Fang [11]. Hence Q has a stochastic representation as $R^2\mathbf{x}$, where R has probability density function $2\pi^{n/2} r^{n-1} f_n(r^2)/\Gamma(n/2)$, $\mathbf{x}$ has Dirichlet distribution $D(k_1/2, \ldots, k_h/2; 1/2)$ and they are independent. The distribution in this central case is denoted by $G_{h+1}(k_1/2, \ldots, k_h/2; 1/2; f_n)$ and studied in Fang [11], see also Anderson and Fang [2] for this type as the distribution of the quadratic form of elliptical variables. If $\alpha = 0$, $\mathbf{z}$ is elliptical with location parameter ξ and density generator $f_{\lambda,k}(x) = \int_{-\infty}^{\lambda} f_n(y_0^2 + x)\, dy_0 / F_1(\lambda)$. The resulted distribution of Q is called the generalized noncentral χ^2 distribution and studied in Fan [10] as the distribution of the quadratic form of elliptical variables. If $\alpha = 0$ and $f_n = \phi_n$, then $Q_i \sim \chi^2_{k_i}(\xi_i' \xi_i)$ $(i = 1, \ldots, h)$ and they are independent. Some basic properties of the noncentral generalized Dirichlet distribution are as follows.

From the construction of Q_i in Proposition 2 it can be seen that if $(Q_1, \ldots, Q_h) \sim NG_h(k_1/2, \ldots, k_h/2; \tau, \lambda; f_n)$, then the vector obtained by adding its components remains of the same type with the parameters k_i and τ added accordingly. For example, $(Q_1, \ldots, Q_{h-2}, Q_{h-1}+Q_h) \sim NG_{h-1}\,(k_1/2, \ldots, k_{h-2}/2; (k_{h-1}+k_h)/2; \tilde{\tau}, \lambda; f_n)$, where $\tilde{\tau}_{i,11} = \tau_{i,11}$, $\tilde{\tau}_{i,12} = \tau_{i,12}$, $\tilde{\tau}_{i,22} = \tau_{i,22}$, $i = 1, \ldots, h-2$; $\tilde{\tau}_{h-1,11} = \tau_{h-1,11} + \tau_{h,11}$, $\tilde{\tau}_{h-1,12} = \tau_{h-1,12} + \tau_{h,12}$, $\tilde{\tau}_{h-1,22} = \tau_{h-1,22} + \tau_{h,22}$.

Since the distribution of the sub-vector of $\mathbf{z}$ in Proposition 2 is of the same type, the distribution of the sub-vector of $(Q_1, \ldots, Q_h)$ is of the same type with the parameters and the density generator modified accordingly. For example, $(Q_1, \ldots, Q_{h-1}) \sim NG_{h-1}$ $(k_1/2, \ldots, k_{h-1}/2; \tilde{\tau}, \tilde{\lambda}; f_{k_0+1})$, where $\tilde{\tau}_{i,11} = \tau_{i,11}$, $\tilde{\tau}_{i,12} = c_1 \tau_{i,12}$, $\tilde{\tau}_{i,22} = c_1^2 \tau_{i,22}$, $\tilde{\lambda} = c_1 \lambda$, $c_1 = (1+\tau_{h,22})^{-1/2}$, $k_0 = \sum_{i=1}^{h-1} k_i$, f_{k_0+1} is the k_0+1-dimensional marginal density generator of f_n. This follows from the fact that $(\mathbf{z}_1', \ldots, \mathbf{z}_{h-1}')' \sim S_{k_0}(\tilde{\xi}, I, \tilde{\lambda}, \tilde{\alpha}; f_{k_0+1})$, where $\tilde{\xi} = (\xi_1', \ldots, \xi_{h-1}')'$, $\tilde{\alpha} = c_1(\alpha_1', \ldots, \alpha_{h-1}')'$, $\tilde{\lambda} = c_1 \lambda$.

3. Moment generating function

In this section, the moment generating functions of the quadratic forms in the skew normal variables are obtained. The proof is based on a basic formula of the integral of the normal distribution function in Azzalini and Dalla Valle [5, p. 719]. The mean and variance of the quadratic form is then obtained by taking derivatives of the moment generation function.

Proposition 3. *Assume* $\mathbf{z} \sim S_k(0, I, \lambda; \alpha; \phi_n)$. *Let* $Q = \mathbf{z}' A \mathbf{z} + 2\mathbf{b}'\mathbf{z} + c$, *where* $A = A'$, $\mathbf{b}$ *and c are constant. Then*

$$E(\exp(tQ)) = \Phi\left(\frac{\lambda + 2\alpha'(I - 2At)^{-1}\mathbf{b}t}{(1 + \alpha'(I - 2At)^{-1}\alpha)^{\frac{1}{2}}}\right) \times \exp(2\mathbf{b}'(I - 2At)^{-1}\mathbf{b}t^2 + ct)|I - 2At|^{-\frac{1}{2}} / \Phi(\lambda / c_0) \tag{8}$$

for small enough t, where c_0 *is given by Definition* 1 *with* $\Omega = I$.

If $\mathbf{z} \sim S_k(\xi, I, \lambda; \alpha; \phi_n)$ and $A_i = A_i'$, substituting $\sum_{i=1}^h t_i A_i$ for tA, $\sum_{i=1}^h t_i A_i \xi$ for $\mathbf{b}t$ and $\xi' \sum_{i=1}^h t_i A_i \xi$ for tc in (8) and using $(I - 2\sum_{i=1}^h A_i t_i)^{-1} 2 \sum_{i=1}^h A_i t_i + I =$

$(I - 2\sum_{i=1}^{h} A_i t_i)^{-1}$, then we obtain

$$E\left(\exp\left(\sum_{i=1}^{h} t_i \mathbf{z}' A_i \mathbf{z}\right)\right) = \Phi\left(\frac{\lambda + 2\alpha'(I - 2\sum_{i=1}^{h} A_i t_i)^{-1}\sum_{i=1}^{h} A_i t_i \xi}{(1+\alpha'(I - 2\sum_{i=1}^{h} A_i t_i)^{-1}\alpha)^{\frac{1}{2}}}\right) \times \exp\left(\xi' \sum_{i=1}^{h} A_i t_i \left(I - 2\sum_{i=1}^{h} A_i t_i\right)^{-1} \xi\right) \times |I - 2\sum_{i=1}^{h} A_i t_i|^{-\frac{1}{2}} / \Phi(\lambda/c_0) \tag{9}$$

for small enough t_i. Let A_i be diagonal with the $\sum_{j=1}^{i-1} k_j + 1$-th to the $\sum_{j=i-1}^{i} k_j$-th diagonal elements being 1, others 0 ($\sum_{j=1}^{0} k_j = 0$), and partition $\mathbf{z}$, ξ and α in the same manner as in Proposition 2, then

$$E\left(\exp\left(\sum_{i=1}^{h} t_i \mathbf{z}_i' \mathbf{z}_i\right)\right) = \Phi\left(\frac{\lambda + \sum_{i=1}^{h} \alpha_i' \xi_i 2t_i (1-2t_i)^{-1}}{(1+\sum_{i=1}^{h} \alpha_i'\alpha_i (1-2t_i)^{-1})^{\frac{1}{2}}}\right) \times \exp\left(\sum_{i=1}^{h} \xi_i'\xi_i t_i (1-2t_i)^{-1}\right) \times \prod_{i=1}^{h} (1-2t_i)^{-\frac{k_i}{2}} / \Phi(\lambda/c_0) \tag{10}$$

for small enough t_i. Eq. (10) gives the moment generating function of $NG_h(k_1/2, \ldots, k_h/2; \tau, \lambda; \phi_n)$ with $\tau_{i,11} = \xi_i'\xi_i$, $\tau_{i,12} = \xi_i'\alpha_i$, $\tau_{i,22} = \alpha_i'\alpha_i$. The moment generating function of one quadratic form of the skew normal variables in the central case ($h = 1$, $\xi = 0$ in our notation) was obtained by Arnold and Beaver [4, p. 31].

Proposition 4. *Assume* $\mathbf{z} \sim S_k(\xi, I, \lambda, \alpha; \phi_n)$ *and A is symmetric. Let* $\zeta_0(x) = \log\{2\Phi(x)\}$, $\zeta_m(x) = d^m/dx^m \zeta_0(x)$ *and* $\mathrm{tr}(A)$ *be the trace of A. Then the first two moments of the quadratic form* $Q = \mathbf{z}'A\mathbf{z}$ *are as follows*:

$$E(Q) = \zeta_1(\lambda/c_0)\left(\frac{2\alpha' A \xi}{c_0} - \frac{\lambda \alpha' A \alpha}{c_0^3}\right) + \xi' A \xi + \mathrm{tr}(A), \tag{11}$$

$$\mathrm{var}(Q) = \zeta_2(\lambda/c_0)\left(\frac{2\alpha' A \xi}{c_0} - \frac{\lambda \alpha' A \alpha}{c_0^3}\right)^2 + \zeta_1(\lambda/c_0)\left(\frac{8\alpha' A^2 \xi}{c_0} - \frac{4\alpha' A \xi \alpha' A \alpha}{c_0^3} - \frac{4\lambda \alpha' A^2 \alpha}{c_0^3} + \frac{3\lambda(\alpha' A \alpha)^2}{c_0^5}\right) + 4\xi' A^2 \xi + 2\,\mathrm{tr}(A^2), \tag{12}$$

where c_0 *is given by Definition* 1 *with* $\Omega = I$.

For the general case that $\mathbf{z}\sim S_k(\xi, \Omega, \lambda, \alpha; \phi_n)$, the moment generating function and moments can be obtained by applying Propositions 3 and 4 to $\Omega^{-1/2}\mathbf{z}$. In the special case that $\lambda = 0$ so that $\zeta_1(0) = (2/\pi)^{1/2}$ and $\zeta_2(0) = -2/\pi$, Proposition 4 leads to the formulas in Genton et al. [13, p. 322]. Note the moments of the quadratic forms can also be obtained from the moments of the underlying random vectors, e.g., $E(\mathbf{z}'A\mathbf{z}) = \mathrm{tr}\{A[E(\mathbf{z}\mathbf{z}')]\}$. The moments of the skew elliptical distribution obtained in Fang [11] can be used in this way to obtain the moments of the quadratic forms in the skew elliptical variables.

4. Cochran's Theorem

In this section we give the necessary and sufficient conditions for the quadratic forms in the skew normal variables to have the noncentral generalized Dirichlet distribution. Let $\mathbf{z}\sim S_k(\xi, I, \lambda, \alpha; \phi_n)$ and $Q_i = \mathbf{z}'A_i\mathbf{z}$. The sufficiency is easy to obtain by using an orthogonal matrix to diagonalize the matrix A_i and holds more generally for the quadratic forms in the skew elliptical variables. The necessity is more difficult to establish, like in the non-skew case. In practice, one often can check the sufficient conditions and conclude that the statistic has a desired distribution. The necessity is of importance in theory. The comments on the Craig's Theorem made in Driscoll and Gundberg [9, p. 65] apply. Theorem 1 is an analogue of the noncentral Cochran's Theorem for the elliptical distributions in Fan [10, Theorem 3.2] and central Cochran's Theorem for the skew elliptical distributions in Fang [11]. Theorem 2 is an extension of the noncentral Cochran's Theorem for the normal distribution, see Rao [19, p. 185], and central Cochran's Theorem for the skew normal distribution in Azzalini and Capitanio [6, Proposition 9]. We first provide a lemma on one quadratic form. One implication of the proof of the lemma is that, given a quadratic form $Q\sim NG_1(k/2; \tau, \lambda; \phi_{k+1})$ in the sense of Definition 2, the parameters are uniquely determined by the distribution. In simpler cases, for example, $Q\sim\chi^2_k(\tau_{11})$, it is easy to identify the parameters. However, with more parameters involved, the work becomes harder.

Lemma 1. *Assume $\mathbf{z}\sim S_k(\xi, I, \lambda, \alpha; \phi_n)$ and $Q = \mathbf{z}'A\mathbf{z}$, where $A = A'$, with $\mathrm{rank}(A) \geqslant 1$, is constant. Then $Q\sim NG_1(k_1/2; \tau, \tilde{\lambda}; \phi_{k_1+1})$ if and only if $\mathrm{rank}(A) = k_1$, $A^2 = A$. In this case, if λ and $\xi'A\alpha$ are not all zeros and $\alpha'A\alpha \neq 0$, then $\tau_{11} = \xi'A\xi$, $\tau_{12} = c_1\xi'A\alpha$, $\tau_{22} = c_1^2\alpha'A\alpha$, $\tilde{\lambda} = c_1\lambda$, where $c_1 = (1 + \alpha'\alpha - \alpha'A\alpha)^{-1/2}$. If $\lambda = 0$ and $\alpha'A\xi = 0$, then the distribution of Q does not depend on $\alpha'A\alpha$ and is $\chi^2_{k_1}(\xi'A\xi)$. If $\alpha'A\alpha = 0$, then the distribution of Q does not depend on λ and is $\chi^2_{k_1}(\xi'A\xi)$.*

Theorem 1. *Assume $\mathbf{z}\sim S_k(\xi, \Omega, \lambda, \alpha; \phi_n)$ and $Q_i = \mathbf{z}'A_i\mathbf{z}$, where $A_i = A_i'$ is constant, $i = 1, \ldots, h$. Let k_i be positive integers such that $\sum_{i=1}^h k_i = k$. Then $(Q_1, \ldots, Q_h)\sim NG_h(k_1/2, \ldots, k_h/2; \tau, \tilde{\lambda}; \phi_n)$ if and only if $\mathrm{rank}(A_i) = k_i$, $A_i\Omega A_i = A_i$, $i = 1, \ldots, h$, $A_i\Omega A_j = 0$, $i \neq j$, $i, j = 1, \ldots, h$. In this case, if λ and $\xi'A_i\Omega\alpha$, $i = 1, \ldots, h$, are not all zeros and $\alpha'\Omega A_i\Omega\alpha$, $i = 1, \ldots, h$, are not all zeros, then $\lambda = \tilde{\lambda}$, $\tau_{i,11} = \xi'A_i\xi$, $\tau_{i,12} = \xi'A_i\Omega\alpha$, $\tau_{i,22} = \alpha'\Omega A_i\Omega\alpha$. Otherwise $Q_i\sim\chi^2_{k_i}(\xi'A_i\xi)$ are independent $(i = 1, \ldots, h)$.*

Theorem 2. *Assume* $\mathbf{z}\sim S_k(\xi, I, \lambda, \alpha; \phi_n)$ *and* $Q_i = \mathbf{z}' A_i \mathbf{z}$, *where* $A_i = A_i'$ *is constant,* $i = 1, \dots, h$, *and* $\sum_{i=1}^h A_i = I$. *Then* $(Q_1, \dots, Q_h)\sim NG_h(k_1/2, \dots, k_h/2; \tau, \tilde{\lambda}; \phi_n)$ *if and only if* $\operatorname{rank}(A_i) = k_i$, $A_i^2 = A_i$, $i = 1, \dots, h$, *if and only if* $\operatorname{rank}(A_i) = k_i$, $\sum_{i=1}^h k_i = k$. *In this case, if* λ *and* $\xi' A_i \alpha$, $i = 1, \dots, h$, *are not all zeros and* $\alpha' A_i \alpha$, $i = 1, \dots, h$, *are not all zeros, then* $\lambda = \tilde{\lambda}$, $\tau_{i,11} = \xi' A_i \xi$, $\tau_{i,12} = \xi' A_i \alpha$, $\tau_{i,22} = \alpha' A_i \alpha$. *If, in addition,* $A_i \alpha \neq 0$ *for at most one i (say* $i = 1$*), then* $Q_1 \sim NG_1(k_1/2; \{\tau_{1,11}, \tau_{1,12}, \tau_{1,22}\}, \lambda; \phi_{k_1+1})$, *and* $Q_i \sim \chi^2_{k_i}(\xi' A_i \xi)$, $i = 2, \dots, h$, *all independent. If* $\lambda = 0$ *and* $\xi' A_i \alpha = 0$, $i = 1, \dots, h$, *or* $\alpha' A_i \alpha = 0$, $i = 1, \dots, h$, *then* $Q_i \sim \chi^2_{k_i}(\xi' A_i \xi)$ *are independent* $(i = 1, \dots, h)$.

5. Application

The quadratic forms arise in many applications. The skew normal distribution can arise from the normal distribution by a truncation mechanism. A stochastic representation of $S_k(\xi, \Omega, \lambda, \alpha; \phi_n)$ is $\mathbf{z} = \xi + \mathbf{x} | x_0 + \lambda/c_0 > 0$, where $(x_0, \mathbf{x}')' \sim N(0, \Sigma)$,

$$\Sigma = \begin{pmatrix} 1 & \delta' \\ \delta & \Omega \end{pmatrix}, \quad \delta = \Omega\alpha/c_0, \quad c_0 = (1 + \alpha'\Omega\alpha)^{1/2},$$

see Fang [11]. The variable x_0 is called hidden variable in Arnold and Beaver [4] or screening variable in Gupta and Brown [14]. The latter authors analyzed a real example in which $\mathbf{x}$ represent the IQ scores of the individuals hired by a company and $\mathbf{z}$ is the observed IQ scores. In this section, we shall give two examples to illustrate the usage of the theory developed for the quadratic forms of the skew normal variables in previous sections.

As the first example, we consider the estimation of the measure of multivariate kurtosis defined by Mardia [17]. Let $\mathbf{z}$ be a p-dimensional random vector with first four moments, then a measure of multivariate kurtosis is $\beta_{2,p} = E(Q^2)$, where $Q = [\mathbf{z} - E(\mathbf{z})]'[\operatorname{cov}(\mathbf{z})]^{-1}[\mathbf{z} - (E\mathbf{z})]$ [17, Eq. (3.5)]. A measure of multivariate skewness $\beta_{1,p}$ is also defined [17, Eq. (2.23)]. Under normality, we have $\beta_{1,p} = 0$ and $\beta_{2,p} = p(p+2)$. The sample analogue of $\beta_{1,p}$ is $b_{1,p}$ and the sample analogue of $\beta_{2,p}$ is $b_{2,p}$ [17, Eqs. (2.23), (3.12)]. A test for multivariate normality that $\beta_{1,p}$ and $\beta_{2,p}$ have the values under normality, using statistics A and B [17, Eqs. (4.1), (4.2)], is formed. The statistic A is a function of $b_{1,p}$ and has χ^2 distribution with $p(p+1)(p+2)/6$ degrees of freedom as asymptotic null distribution. The statistic B is a function of $b_{2,p}$ and has $N(0, 1)$ as asymptotic null distribution. Suppose now $\mathbf{z}\sim S_p(\xi, \Omega, \lambda, \alpha; \phi_{p+1})$. Then $\mathbf{z} - E(\mathbf{z}) \sim S_p(-\zeta_1(\lambda/c_0)\Omega\alpha/c_0, \Omega, \lambda, \alpha; \phi_{p+1})$, see Arnold and Beaver [4, Eq. (4.15)] for the calculation of $E(\mathbf{z})$. Hence $\beta_{2,p}$ can be calculated by Proposition 4 with suitable parameters. We omit the derivation of $\beta_{1,p}$ since Proposition 4 does not provide full basis for the calculation. The data set reported by Cook and Weisberg [8] for 202 athletes at the Australian Institute of Sport was analyzed by Arnold and Beaver [4]. They obtained the maximum likelihood estimates of the parameters, assuming the person's height and weight have joint distribution $S_2(\mu, \Sigma, \lambda_0, \Sigma^{-1/2}(\lambda_1, \lambda_2)'; \phi_3)$ in the notation here. Denote the diagonal elements of $\Sigma^{-1/2}$ by θ_1, θ_3 and its off-diagonal elements by θ_2. The estimates are obtained for three cases: the model with 5 parameters, assuming $\lambda_0 = 0$, $\lambda_1 = 0$, $\lambda_2 = 0$; the model with 7 parameters, assuming $\lambda_0 = 0$; the model with 8 parameters. Substituting these estimates into $\beta_{2,p}$ obtained above with $p = 2$, we obtain its estimate $\hat{\beta}_{2,p}$. In the first case the distribution of $\mathbf{z}$ is normal and $\beta_{2,p} = 8$, which does not depend on the parameters. We obtain $\hat{\beta}_{2,p} = 8.5449$ in the second case, $\hat{\beta}_{2,p} = 9.1655$ in the third case. Using the data of the height and weight of the 202 athletes, we obtain $b_{1,p} = 1.6878$, $b_{2,p} = 10.8103$, $A = 56.8229$ and $B = 4.9928$. Since $A > 9.4877$, the critical value of χ^2_4 distribution, and $B > 1.96$, the hypotheses $\beta_{1,p} = 0$ and $\beta_{2,p} = p(p+2)$ are both rejected at level 0.05. We conclude the data cannot be regarded as a sample from a normal population. Moreover, $\hat{\beta}_{2,p}$ in the model with more parameters is closer to the observed $b_{2,p}$, indicating the model provides a better fit. This finding coincides with that obtained by likelihood ratio test in Arnold and Beaver [4].

In the second example, we shall investigate the least-squares estimator when the error differs from the normal by having a skew normal distribution. Gupta and Brown [14] obtained maximum likelihood estimates of the parameters from independent, identically distributed univariate skew normal variables. We here consider a different model where the sample $(z_1, \ldots, z_k)$ are jointly skew normal

$$\mathbf{z} = \xi + \varepsilon, \tag{13}$$

where $\varepsilon \sim S_k(0, I\sigma^2, \lambda, \alpha; \phi_n)$, $\xi = X\beta$, X is $k \times m$ with rank $m < k$, β is $m \times 1$. If $\alpha = 0$, then we recover $\mathbf{z}$ as normal with location parameter ξ and covariance $I\sigma^2$. The least-squares estimator of β is $\hat{\beta} = (X'X)^{-1}X'\mathbf{z}$, the estimator of ξ is $\hat{\xi} = P\mathbf{z}$ and the sum of residual squares is $R_0^2 = (\mathbf{z} - \hat{\xi})'(\mathbf{z} - \hat{\xi}) = \mathbf{z}'(I - P)\mathbf{z}$, where $P = X(X'X)^{-1}X'$ is the projector matrix on the range of X, $\mathcal{L}(X)$. The estimator $\hat{\beta}$ is unbiased for β and $\hat{\sigma}^2 = R_0^2/(k-m)$ is an unbiased estimator of σ^2, $\hat{\beta}$ and $\hat{\sigma}^2 = R_0^2/(k-m)$ are independent.

Consider the general case that α is arbitrary. Let $\mathbf{y} = \mathbf{z}/\sigma$, $A_1 = P$, $A_2 = I - P$ and $Q_i = \mathbf{y}'A_i\mathbf{y}$, then $\hat{\xi}'\hat{\xi} = \sigma^2 Q_1$, $R_0^2 = \sigma^2 Q_2$. Since $\mathbf{y} \sim S_k(\xi/\sigma, I, \lambda, \sigma\alpha; \phi_n)$, by Definition 2, $(Q_1, Q_2) \sim NG_2(m/2, (k-m)/2; \tau, \lambda; \phi_n)$, where $\tau_{1,11} = \xi' P \xi \sigma^{-2} = \xi'\xi\sigma^{-2}$, $\tau_{1,12} = \xi'\alpha$, $\tau_{1,22} = \alpha' P \alpha \sigma^2$, $\tau_{2,11} = \xi'(I-P)\xi\sigma^{-2} = 0$, $\tau_{2,12} = \xi'(I-P)\alpha = 0$, $\tau_{2,22} = \alpha'(I-P)\alpha\sigma^2$. Marginally, $Q_1 \sim NG_1(m/2; \{\tau_{1,11}, c_1\tau_{1,12}, c_1^2\tau_{1,22}\}, c_1\lambda; \phi_{m+1})$, $Q_2 \sim NG_1((k-m)/2; \{0, 0, c_2\tau_{2,22}\}, c_2\lambda; \phi_{k-m+1})$, where $c_1 = (1 + \alpha'(I-P)\alpha\sigma^2)^{-1/2}$, $c_2 = (1 + \alpha' P \alpha \sigma^2)^{-1/2}$. Let $X = \Gamma_1 B$, where Γ_1 is $k \times m$, $\Gamma_1'\Gamma_1 = I$, B is the positive definite square root of $X'X$. Then calculation shows $\hat{\beta} = B^{-1}\Gamma_1'\mathbf{z} \sim S_m(\beta, (X'X)^{-1}\sigma^2, c_1\lambda, c_1 X'\alpha; \phi_{m+1})$. By Arnold and Beaver [4],

$$E(\mathbf{z} - \xi) = \zeta_1(\lambda/c_0)\sigma^2 c_0^{-1}\alpha, \tag{14}$$

$$E(\mathbf{z} - \xi)(\mathbf{z} - \xi)' = I\sigma^2 - \zeta_1(\lambda/c_0)\lambda\sigma^4 c_0^{-3}\alpha\alpha', \tag{15}$$

where ζ_1 is defined in Proposition 4, $c_0 = (1 + \sigma^2\alpha'\alpha)^{1/2}$. Hence

$$\begin{aligned} E(\hat{\beta}) &= (X'X)^{-1}X'\xi + (X'X)^{-1}X'E(\mathbf{z} - \xi) \\ &= \beta + \zeta_1(\lambda/c_0)\sigma^2 c_0^{-1}(X'X)^{-1}X'\alpha. \end{aligned} \tag{16}$$

$$E(\hat{\xi}) = \xi + \zeta_1(\lambda/c_0)\sigma^2 c_0^{-1} P\alpha, \tag{17}$$

$$E(R_0^2) = \operatorname{tr}\{(I-P)E(\mathbf{z}-\xi)(\mathbf{z}-\xi)'(I-P)\}$$
$$= (k-m)\sigma^2 - \zeta_1(\lambda/c_0)\lambda\sigma^4 c_0^{-3}\alpha'(I-P)\alpha. \tag{18}$$

We consider some special cases as follows. If $\alpha \in \mathcal{L}(X)$, for example, the components of α are equal and X contains a column with equal components. By (14) and (17), the mean of the vector of residuals, $\mathbf{z}-\hat{\xi}$, is zero. Since $\operatorname{cov}(\mathbf{z}-\xi, \hat{\beta}) = (I-P)\operatorname{var}(\mathbf{z}-\xi)X(X'X)^{-1}$, by (14) and (15) $\mathbf{z}-\hat{\xi}$ and $\hat{\beta}$ are uncorrelated, $\mathbf{z}-\hat{\xi}$ and $\hat{\xi}$ are uncorrelated. However, $\hat{\beta}$ is not necessarily an unbiased estimator of β. Since $A_2\alpha = 0$ and $\tau_{2,22} = 0$, by Theorem 2, $\hat{\xi}'\hat{\xi} \sim \sigma^2\, NG_1(m/2; \{\tau_{1,11}, \tau_{1,12}, \tau_{1,22}\}, \lambda; \phi_{m+1})$, $R_0^2 \sim \sigma^2\chi^2_{k-m}$ and they are independent. An unbiased estimator of σ^2 is $\hat{\sigma}^2 = (k-m)^{-1}R_0^2$, the same as in the case $\alpha = 0$. If $\alpha \in \mathcal{L}(X)^{\perp}$, then $\hat{\beta} \sim N(\beta, (X'X)^{-1}\sigma^2)$ and is an unbiased estimator of β. Since $A_1\alpha = 0$, $\tau_{1,22} = 0$, $\hat{\xi}'\hat{\xi} \sim \chi^2_m(\xi'\xi)$, $R_0^2 \sim \sigma^2 NG_1((k-m)/2; \{0, 0, \tau_{2,22}\}, \lambda; \phi_{k-m+1})$, and they are independent. In this case, as in the case that $\alpha \neq 0$, it seems not feasible to obtain an unbiased estimator of σ^2 from R_0^2 by (18). In general, the optimal properties of the conventional least-squares method for the normal distributions are not maintained if the data in fact arise with an additional truncation.

For numerical illustration, we shall calculate the true level of the confidence interval based on R_0^2 for σ^2. Suppose $[R_0^2/u_2, R_0^2/u_1]$ is the confidence interval with specified level under the usual normality assumption. The true level when the data actually arise from (13) is $P(u_1 < Q_2 < u_2)$. Table 1 presents the true level for various skew parameters with $k = 10$, $m = 1$ and the level 0.95. We have $u_1 = 2.7004$, $u_2 = 19.0228$. Since the distribution of Q_2 depends on the skew parameters only though $c_2\lambda$ and $c_2\tau_{2,22}$, we present the probability that $NG_1((k-m)/2, \{0, 0, \tau_{22}\}, \lambda; \phi_{k-m+1})$ lies in the interval $[u_1, u_2]$ for various combination of λ and τ_{22}. Proposition 2 is used for the calculation. To apply formula (7) for the skew normal generator with $h = 1$ and $k_1 \geqslant 3$, an integration on the three-dimensional space is needed. In our case some of the parameters vanish so that the dimension can be reduced further. Starting from (22) in the proof of Proposition 1 in Section 6 with $l_1 = 0$, $l_2 = 0$, $\tau_{11} = 0$, $\tau_{12} = 0$ and using the formula in Fang et al. [12, p. 23], we reduce the integral for variables $y_2, \ldots, y_k$ to an integral on the one-dimensional space if $k \geqslant 2$. Denote the new variable by y_2, the distribution function of $NG_1(k/2, \{0, 0, \tau_{22}\}, \lambda; \phi_{k+1})$ given by (22) is equal to

$$\frac{\pi^{(k-1)/2}}{\Gamma(\frac{k-1}{2})}\int_{R^2} I_{(0,x)}(\mathbf{y}'\mathbf{y})\left[\int I_{(-\infty,\lambda+\tau_{22}^{\frac{1}{2}}y_1)}(y_0)|y_2|^{k-2} f_n(y_0^2+\mathbf{y}'\mathbf{y})\,dy_0\right]d\mathbf{y}/F_1(\frac{\lambda}{c_0}).$$

If $f_n = \phi_n$, then the above equation is equal to

$$\frac{2^{1-k/2}\pi^{1/2}}{\Gamma(\frac{k-1}{2})}\int_{R^2} I_{(0,x)}(\mathbf{y}'\mathbf{y})\Phi(\lambda+\tau_{22}^{\frac{1}{2}}y_1)|y_2|^{k-2}\phi(y_1)\phi(y_2)\,d\mathbf{y}/\Phi(\frac{\lambda}{c_0}).$$

Table 1 shows if $\lambda > 0$, then the true level is larger than its nominal value for small $c_2\tau_{2,22}$ and smaller than its nominal value for large $c_2\tau_{2,22}$. For fixed $\lambda (> 0)$, the true level increases with $c_2\tau_{2,22}$ in the neighborhood of zero and decreases when $c_2\tau_{2,22}$ exceeds certain value. The opposite holds if $\lambda < 0$. For the range of the parameters used in the calculation,

Table 1
True level of the confidence interval for σ^2

$c_2\lambda \backslash c_2\tau_{2,22}$	0.1	0.3	0.5	1	2	3	4
−2.0000	0.9466	0.9427	0.9409	0.9423	0.9468	0.9498	0.9516
−1.0000	0.9491	0.9484	0.9484	0.9493	0.9509	0.9517	0.9521
−0.5000	0.9497	0.9495	0.9495	0.9499	0.9506	0.9509	0.9511
0.5000	0.9502	0.9503	0.9502	0.9500	0.9496	0.9494	0.9492
1.0000	0.9502	0.9504	0.9504	0.9502	0.9497	0.9492	0.9490
2.0000	0.9501	0.9503	0.9505	0.9507	0.9505	0.9500	0.9496

Note: The entries are $P(u_1 < Q_2 < u_2)$, with $k = 10$, $m = 1$. The nominal level is 0.95.

the departure of the true level from the nominal level is not severe. If $\alpha \in \mathcal{L}(X)^{\perp}$, then $c_2 = 1$ and $\tau_{2,22} = \sigma^2\alpha'\alpha$ increases with σ^2 and $\alpha'\alpha$. Note if $\alpha \in \mathcal{L}(X)$ or $\lambda = 0$, then the distribution of Q_2 is invariant for model (13) and the true level is equal to the nominal level by Lemma 1.

6. Proofs

Proof of Proposition 1. Using the probability density function of $\mathbf{z}$, we obtain

$$\mathrm{pr}(Q < x) = \int_{R^k} I_{(0,x)}(\mathbf{z}'\mathbf{z}) \left[\int I_{(-\infty,\lambda+\alpha'\mathbf{z}-\alpha'\xi)}(y_0) \times f_n(y_0^2 + \mathbf{z}'\mathbf{z} - 2\xi'\mathbf{z} + \xi'\xi)\, dy_0 \right] d\mathbf{z}/F_1(\frac{\lambda}{c_0}). \tag{19}$$

If $k = 1$, (19) is equal to

$$\int I_{(0,x)}(y^2) \left[\int I_{(-\infty,\lambda-\alpha y-\alpha\xi)}(y_0) f_n(y_0^2 + y^2 + 2\xi y + \xi^2)\, dy_0 \right] dy/F_1(\frac{\lambda}{c_0}). \tag{20}$$

If $\alpha > 0$, then $l_1 = \xi\alpha/|\alpha| = \xi$ and $\tau_{22}^{1/2} = \alpha$. If $\alpha < 0$, then $l_1 = -\xi$ and $\tau_{22}^{1/2} = -\alpha$. If $\alpha = 0$ and $\xi \geqslant 0$, then $l_1 = \xi$ and $\tau_{22}^{1/2} = \alpha$. If $\alpha = 0$ and $\xi < 0$, then $l_1 = -\xi$ and $\tau_{22}^{1/2} = -\alpha$. Substitute these expressions into (19) for the cases $l_1 = \xi$ and (20) for $l_1 = -\xi$ we obtain the distribution function of Q as

$$\mathrm{pr}(Q < x) = \int I_{(0,x)}(y^2) \left[\int I_{(-\infty,\lambda+\tau_{22}^{1/2} y-\tau_{12})}(y_0) \times f_n(y_0^2 + y^2 - 2l_1 y + \tau_{11})\, dy_0 \right] dy/F_1(\lambda/c_0). \tag{21}$$

This leads to (4). If $k \geqslant 2$, we consider four cases. In the case that $\alpha \neq 0$ and ξ is not proportional to α so that $\tau_{22} \neq 0$ and $l_2 \neq 0$, let Γ be an orthogonal matrix with $\alpha/\|\alpha\|$ and $(\xi - l_1\alpha/\|\alpha\|)/l_2$ as its first two columns. Make transformation $\mathbf{y} = \Gamma'\mathbf{z}$ in (19) to obtain

$$\mathrm{pr}(Q < x) = \int_{R^k} I_{(0,x)}(\mathbf{y}'\mathbf{y}) \left[\int I_{(-\infty,\lambda+\tau_{22}^{\frac{1}{2}} y_1-\tau_{12})}(y_0) \times f_n(y_0^2 + \mathbf{y}'\mathbf{y} - 2(l_2 y_2 + l_1 y_1) + \tau_{11})\, dy_0 \right] d\mathbf{y}/F_1(\frac{\lambda}{c_0}). \tag{22}$$

In the case that $\alpha \neq 0$ and ξ is proportional to α so that $\tau_{22} \neq 0$ and $l_2 = 0$, let Γ be an orthogonal matrix with $\alpha/\|\alpha\|$ as its first column. Make transformation $\mathbf{y} = \Gamma'\mathbf{z}$ in (19) to obtain (22) with $l_2 = 0$. In the case $\alpha = 0$ and $\xi \neq 0$ so that $\tau_{22} = 0$, $\tau_{12} = 0$, $\tau_{11} \neq 0$, let Γ be an orthogonal matrix with $\xi/\|\xi\|$ as its second column. Make transformation $\mathbf{y} = \Gamma'\mathbf{z}$ in (19) to obtain (22) with $l_1 = 0$. In the case that $\alpha = 0$ and $\xi = 0$ so that $\tau_{22} = 0$, $\tau_{12} = 0$ and $\tau_{11} = 0$ (19) is identical to (22) with $l_1 = 0$ and $l_2 = 0$. These expressions give the distribution function for $k = 2$ in a unified form (22). By successive transformations $y_1 = r\cos\theta$, $y_2 = r\sin\theta$ and then $r^2 = s$, (22) with $k = 2$ is equal to

$$\frac{1}{2F_1(\lambda/c_0)} \int_{\substack{0<\theta<2\pi \\ s>0}} I_{(0,x)}(s) \left[\int I_{(-\infty,\lambda+\tau_{22}^{\frac{1}{2}} s^{\frac{1}{2}} \cos\theta-\tau_{12})}(y_0) \times f_n(y_0^2 + s - 2s^{\frac{1}{2}}(l_2 \sin\theta + l_1 \cos\theta) + \tau_{11})\, dy_0 \right] ds\, d\theta,$$

which leads to (5). If $k \geqslant 3$, by a formula in Fang et al. [12, p. 23], the dimension of y can be reduced and (22) is equal to

$$\mathrm{pr}(Q < x) = c(k) \int_{R^3} I_{(0,x)}(\mathbf{y}'\mathbf{y})|y_3|^{k-3} \left[\int I_{(-\infty,\lambda+\tau_{22}^{1/2} y_1-\tau_{12})}(y_0) \times f_n(y_0^2 + \mathbf{y}'\mathbf{y} - 2(l_2 y_2 + l_1 y_1) + \tau_{11})\, dy_0 \right] d\mathbf{y}/F_1(\lambda/c_0), \tag{23}$$

which can be expressed, by making successive transformations $y_1 = r\cos\theta_1$, $y_2 = r\sin\theta_1\cos\theta_2$, $y_3 = r\sin\theta_1\sin\theta_2$ and then $r^2 = s$, as

$$\frac{c(k)}{2F_1(\lambda/c_0)} \int_{\substack{0<\theta_1<\pi \\ 0<\theta_2<2\pi \\ s>0}} s^{\frac{k}{2}-1} I_{(0,x)}(s)(\sin\theta_1)^{k-2}|\sin\theta_2|^{k-3} \left[\int I_{(-\infty,\lambda+\tau_{22}^{\frac{1}{2}} s^{\frac{1}{2}} \cos\theta_1-\tau_{12})}(y_0) \times f_n(y_0^2 + s - 2s^{\frac{1}{2}}(l_2\sin\theta_1\cos\theta_2 + l_1\cos\theta_1) + \tau_{11})\, dy_0 \right] ds\, d\theta_1\, d\theta_2.$$

This leads to (6). □

Proof of Proposition 2. Using the probability density function of $\mathbf{z}$, we obtain

$$\begin{aligned}
&\mathrm{pr}(Q_1 < x_1, \ldots, Q_h < x_h) \\
&\quad = \int_{\prod_{i=1}^h R^{k_i}} \prod_{i=1}^h I_{(0,x_i)}(\mathbf{z}_i'\mathbf{z}_i) \left[\int I_{(-\infty,\lambda+\sum_{i=1}^h \alpha_i'\mathbf{z}_i - \sum_{i=1}^h \alpha_i'\xi_i)}(y_0) \right. \\
&\qquad \left. \times f_n\left(y_0^2 + \sum_{i=1}^h \mathbf{z}_i'\mathbf{z}_i - 2\sum_{i=1}^h \xi_i'\mathbf{z}_i + \sum_{i=1}^h \xi_i'\xi_i \right) dy_0 \right] \prod_{i=1}^h d\mathbf{z}_i / F_1(\lambda/c_0). \quad (24)
\end{aligned}$$

Making transformations on $\mathbf{z}_i$ in (24) according to $k_i = 1$, $k_i = 2$ or $k_i \geqslant 3$ as in the proof of Proposition 1 simultaneously we obtain (7), see (21)–(23). $\square$

Proof of Proposition 3. Using the probability density function of $\mathbf{z}$ with $f_n = \phi_n$, we obtain

$$\begin{aligned}
&E\exp(tQ) \\
&\quad = \int \exp(t(\mathbf{z}'A\mathbf{z} + 2\mathbf{b}'\mathbf{z} + c)) \left\{ \int_{-\infty}^{\lambda+\alpha'\mathbf{z}} \exp(-(y_0^2 + \mathbf{z}'\mathbf{z})/2)\, dy_0 \right\} d\mathbf{z} \\
&\qquad \times (2\pi)^{-n/2}/\Phi(\lambda/c_0) \\
&\quad = \int \Phi(\lambda + 2\alpha'(I-2At)^{-1}bt + \alpha'(\mathbf{z} - 2(I-2At)^{-1}\mathbf{b}t)) \\
&\qquad \times \exp(-\frac{1}{2}(\mathbf{z} - 2(I-2At)^{-1}\mathbf{b}t)'(I-2At)(\mathbf{z} - 2(I-2At)^{-1}\mathbf{b}t))(2\pi)^{-k/2} \\
&\qquad \times |I-2At|^{1/2}\, d\mathbf{z} \exp(2\mathbf{b}'(I-2At)^{-1}\mathbf{b}t^2 + ct)|I-2At|^{-1/2}/\Phi(\lambda/c_0),
\end{aligned}$$

which is equal to (8) by Azzalini and Dalla Valle [5, Proposition 4]. $\square$

Proof of Proposition 4. By (9), let the moment generating function be denoted by

$$M(t) = \Phi(I_1)\exp(I_2)I_3/\Phi(\lambda/c_0), \quad (25)$$

where

$$\begin{aligned}
I_1 &= (\lambda + 2\alpha'(I-2At)^{-1}A\xi t)(1 + \alpha'(I-2At)^{-1}\alpha)^{-\frac{1}{2}} = I_{11}I_{12}, \\
I_2 &= \xi' A(I-2At)^{-1}\xi t, \quad I_3 = |I-2At|^{-\frac{1}{2}}, \quad c_0 = (1+\alpha'\alpha)^{1/2}.
\end{aligned}$$

Then

$$\frac{d}{dt}(\log\ M(t)) = \zeta_1(I_1)\frac{dI_1}{dt} + \frac{dI_2}{dt} + \frac{d(\log(I_3))}{dt}, \quad (26)$$

$$\frac{d^2}{dt^2}(\log M(t)) = \zeta_2(I_1)\left(\frac{dI_1}{dt}\right)^2 + \zeta_1(I_1)\frac{d^2I_1}{dt^2} + \frac{d^2I_2}{dt^2} + \frac{d^2(\log(I_3))}{dt^2}, \quad (27)$$

which leads to $E(Q)$ and $\mathrm{var}(Q)$ by letting $t = 0$. Suppose A is diagonal with diagonal elements a_i. Then by calculation,

$$
\begin{aligned}
\frac{dI_{11}}{dt} &= \sum \frac{2a_i\alpha_i\xi_i}{(1-2a_it)^2}, \quad \frac{d^2I_{11}}{dt^2} = \sum \frac{8a_i^2\alpha_i\xi_i}{(1-2a_it)^3},\\
\frac{dI_{12}}{dt} &= -I_{12}^3 \sum \frac{a_i\alpha_i^2}{(1-2a_it)^2},\\
\frac{d^2I_{12}}{dt^2} &= 3I_{12}^5 \left(\sum \frac{a_i\alpha_i^2}{(1-2a_it)^2}\right)^2 - I_{12}^3 \sum \frac{4a_i^2\alpha_i^2}{(1-2a_it)^3},\\
\frac{dI_1}{dt} &= \left. \frac{dI_{11}}{dt} I_{12} + I_{11}\frac{dI_{12}}{dt}\right|_{t=0} = \frac{2\alpha' A\xi}{c_0} - \frac{\lambda\alpha' A\alpha}{c_0^3},\\
\frac{d^2I_1}{dt^2} &= \left. \frac{d^2I_{11}}{dt^2} I_{12} + 2\frac{dI_{11}}{dt}\,\frac{dI_{12}}{dt} + I_{11}\frac{d^2I_{12}}{dt^2}\right|_{t=0}\\
&= \frac{8\alpha' A^2\xi}{c_0} - \frac{4\alpha' A\xi\alpha' A\alpha}{c_0^3} + \lambda(\frac{3(\alpha' A\alpha)^2}{c_0^5} - \frac{4\alpha' A^2\alpha}{c_0^3}),\\
\frac{dI_2}{dt} &= \left. \sum \frac{a_i\xi_i^2}{(1-2a_it)^2}\right|_{t=0} = \xi' A\xi,\\
\frac{d^2I_2}{dt^2} &= \left. \sum \frac{4a_i^2\xi_i^2}{(1-2a_it)^3}\right|_{t=0} = 4\xi' A^2\xi,\\
\frac{d\log(I_3)}{dt} &= \left. \sum \frac{a_i}{1-2a_it}\right|_{t=0} = \mathrm{tr}(A),\\
\frac{d^2\log(I_3)}{dt^2} &= \left. \sum \frac{2a_i^2}{(1-2a_it)^2}\right|_{t=0} = 2\,\mathrm{tr}(A^2).
\end{aligned}
$$

Substituting these values into (26) and (27) we obtain (11) and (12). In general, if A is symmetric, let Γ be an orthogonal matrix such that $\Gamma' A\Gamma$ is diagonal and $\mathbf{y} = \Gamma'\mathbf{z}$. Then $\mathbf{y} \sim S_k(\Gamma'\xi, I, \lambda, \Gamma'\alpha; \phi_n)$, $\mathbf{y}'(\Gamma' A\Gamma)\mathbf{y} = \mathbf{z}' A\mathbf{z}$, $(\Gamma'\xi)'(\Gamma' A\Gamma)(\Gamma'\xi) = \xi' A\xi$, etc. Applying the established formulas (11) and (12) to the quadratic form of $\mathbf{y}$, we obtain those for $\mathbf{z}$. $\square$

Proof of Lemma 1. *Sufficiency*: By assumption, there exists an orthogonal matrix Γ such that $\Gamma' A\Gamma$ is diagonal with the first k_1 diagonal elements being 1 and others 0. Partition Γ as (Γ_1, Γ_2), where Γ_1 is $k \times k_1$. Then $A = \Gamma_1\Gamma_1'$. Let $\mathbf{y} = \Gamma'\mathbf{z}$. Then $\mathbf{y} \sim S_k(\Gamma'\xi, I, \lambda, \Gamma'\alpha; \phi_n)$. Let $\mathbf{y}_1$ be the sub-vector of $\mathbf{y}$ consisting of its first k_1 components. Then $\mathbf{y}_1 \sim S_{k_1}(\Gamma_1'\xi, I, c_1\lambda, c_1\Gamma_1'\alpha; \phi_{k_1+1})$, where $c_1 = (1 + (\Gamma_2'\alpha)'\Gamma_2'\alpha)^{-1/2} = (1 + \alpha'\alpha - \alpha' A\alpha)^{-1/2}$, and by Definition 2 $\mathbf{z}' A\mathbf{z} = \mathbf{y}_1'\mathbf{y}_1 \sim NG_1(k_1/2; \tau, \tilde{\lambda}; \phi_{k_1+1})$, where $\tau_{11} = \xi'\Gamma_1\Gamma_1'\xi = \xi' A\xi$, $\tau_{12} = \xi'\Gamma_1 c_1\Gamma_1'\alpha = c_1\xi' A\alpha$, $\tau_{22} = c_1\alpha'\Gamma_1 c_1\Gamma_1'\alpha = c_1^2\alpha' A\alpha$, $\tilde{\lambda} = c_1\lambda$.

Necessity: The moment generating function of $\mathbf{z}' A\mathbf{z}$ is (25) in the proof of Proposition 4. The moment generating function of $NG_1(k_1/2; \tau, \tilde{\lambda}; \phi_{k_1+1})$ is, by (10) and

Definition 2,

$$\Phi(\tilde{I}_1)\exp(\tilde{I}_2)\tilde{I}_3/\Phi(\tilde{\lambda}/\tilde{c}_0), \tag{28}$$

where

$$\begin{aligned}\tilde{I}_1 &= (\tilde{\lambda}+2\tau_{12}t(1-2t)^{-1})(1+\tau_{22}(1-2t)^{-1})^{-\frac{1}{2}},\\ \tilde{I}_2 &= \tau_{11}t(1-2t)^{-1}, \quad \tilde{I}_3=(1-2t)^{-\frac{k_1}{2}},\\ \tilde{c}_0 &= (1+\tau_{22})^{\frac{1}{2}}.\end{aligned}$$

Following the argument of Laha [15], see also Driscoll and Gundberg [9, p. 67], let

$$h(t)=\left\{\frac{\Phi(I_1)\exp(I_2)\Phi(\frac{\tilde{\lambda}}{\tilde{c}_0})}{\Phi(\tilde{I}_1)\exp(\tilde{I}_2)\Phi(\frac{\lambda}{c_0})}\right\}^2-\frac{I_3^{-2}}{\tilde{I}_3^{-2}}, \tag{29}$$

where t is now a complex variable and Φ is expressed by confluent hypergeometric function ${}_1F_1$,

$$\Phi(I_1)=\frac{1}{2}+\frac{I_1}{(2\pi)^{1/2}}{}_1F_1(\frac{1}{2},\frac{3}{2},-\frac{I_1^2}{2}),$$

see Abramowitz and Stegun [1, p. 298, 7.1.21, 7.1.22]. Then $h(t)=0$ for t real and small enough by the equality of (25) and (28). As a function of complex variable, $h(t)$ is analytic so that $h(t)=0$ for all complex t with $|I-2At|\neq 0$ and $t\neq 1/2$. Since the first term of (29) has no zeros, its second term has none. This implies the second term, as the ratio of two polynomials, must be constant. Thus $|I-2At|=(1-2t)^{k_1}$, implying $\operatorname{rank}(A)=k_1$ and the non-zero latent roots of A are all 1 so that $A^2=A$. By the proof of sufficiency, the moment generating function of $\mathbf{z}'A\mathbf{z}$ is $\Phi(I_1)\exp(I_2)(1-2t)^{-k_1/2}/\Phi(c_1\lambda(1+c_1^2\alpha'A\alpha)^{-1/2})$, where

$$\begin{aligned}I_1 &= (c_1\lambda+2c_1\alpha'A\xi t(1-2t)^{-1})(1+c_1^2\alpha'A\alpha(1-2t)^{-1})^{-\frac{1}{2}},\\ I_2 &= \xi'A\xi t(1-2t)^{-1}.\end{aligned}$$

Since $c_1\lambda(1+c_1^2\alpha'A\alpha)^{-1/2}=\lambda/c_0$, we then have

$$\Phi(I_1)\exp(I_2)/\Phi(\lambda/c_0)=\Phi(\tilde{I}_1)\exp(\tilde{I}_2)/\Phi(\tilde{\lambda}/\tilde{c}_0) \tag{30}$$

by the equality of the two moment generating functions.

We shall prove that (30) implies the relation of the parameters stated in the lemma. For notation simplicity we replace $c_1\alpha$ by α, $c_1\lambda$ by λ and assume $c_1=1$. This is equivalent to the assumption that A is of full rank. We shall first prove that $\alpha'\xi$ and τ_{12} have the same sign. Then we deduce $\xi'\xi=\tau_{11}$. This leads to the relation that $I_1=\tilde{I}_1$, with which we identify the rest parameters.

Let $y=2t(1-2t)^{-1}$ in I_j and use the same notation I_j to denote the resulted functions. Let $I_j^{(1)}=d/dy(I_j)$. Then

$$\begin{aligned}I_1 &= (\lambda+\alpha'\xi y)(1+\alpha'\alpha+\alpha'\alpha y)^{-1/2},\\ I_1^{(1)} &= [2(1+\alpha'\alpha)\alpha'\xi-\lambda\alpha'\alpha+\alpha'\alpha\alpha'\xi y](1+\alpha'\alpha+\alpha'\alpha y)^{-3/2}/2,\\ I_2 &= \xi'\xi y/2, \quad I_2^{(1)}=\xi'\xi/2.\end{aligned}$$

Let

$$B = \frac{d}{dy}\ \log(\Phi(I_1)\exp(I_2)/\Phi(\lambda/c_0)) = \zeta_1(I_1)I_1^{(1)} + I_2^{(1)},$$

where ζ_1 is defined in Proposition 4. Define $\tilde{B}$ similarly. Then $B = \tilde{B}$.

Let $y \to +\infty$. If $\alpha'\xi > 0$, then $I_1 \to +\infty$, $\zeta_1(I_1) \to 0$, $I_1^{(1)} \to 0$. If $\alpha'\xi = 0$ and $\alpha'\alpha > 0$, then $I_1 \to 0$, $\zeta_1(I_1) \to (2/\pi)^{1/2}$, $I_1^{(1)} \to 0$. If $\alpha'\xi = 0$ and $\alpha'\alpha = 0$, then $I_1 = \lambda$, $I_1^{(1)} = 0$. Hence if $\alpha'\xi \geqslant 0$, then $B \to \xi'\xi/2$. If $\alpha'\xi < 0$, then $I_1 \to -\infty$, by a formula in Abramowitz and Stegun [1, p. 298, 7.1.23],

$$\begin{aligned} \Phi(I_1) &= 1 - \Phi(-I_1) \\ &= -I_1^{-1}(2\pi)^{-1/2}\exp(-I_1^2/2)(1 - I_1^{-2} + o(I_1^{-2})) \\ &= -I_1^{-1}\phi(I_1)(1 + o(I_1^{-1})). \end{aligned} \tag{31}$$

Hence

$$\zeta_1(I_1) = -I_1(1 + o(I_1^{-1})).$$

It follows that $B \to -(\alpha'\xi)^2/(2\alpha'\alpha) + \xi'\xi/2$. By (30),

$$\frac{\Phi(I_1)\Phi(\tilde{\lambda}/\tilde{c_0})}{\Phi(\lambda/c_0)} = \Phi(\tilde{I}_1)\exp(\tilde{I}_2 - I_2). \tag{32}$$

Assume $\alpha'\xi \geqslant 0$ and $\tau_{12} < 0$, then $\Phi(I_1)$ has non-zero lower bound and the left-hand side of (32) has non-zero finite limit. Since $\tilde{I}_1 \to -\infty$, the right-hand side of (32) is equivalent to $\tilde{I}_1^{-1}\exp(-\tilde{I}_1^2/2 + \tilde{I}_2 - I_2)(2\pi)^{-1/2}$, with limit 0 or ∞. This contradiction shows that if $\alpha'\xi \geqslant 0$ then $\tau_{12} \geqslant 0$ and vice versa. If $\alpha'\xi \geqslant 0$ and $\tau_{12} \geqslant 0$, then by the equality of B and $\tilde{B}$, we obtain $\xi'\xi = \tau_{11}$. If $\alpha'\xi < 0$ and $\tau_{12} < 0$, then

$$-(\alpha'\xi)^2/\alpha'\alpha + \xi'\xi = -\tau_{12}^2/\tau_{22} + \tau_{11}. \tag{33}$$

By (31) and (30), as $y \to \infty$,

$$\frac{\Phi(\lambda/c_0)}{\Phi(\tilde{\lambda}/\tilde{c}_0)} \sim \frac{\tilde{I}_1}{I_1}\exp\left(-\frac{I_1^2 - \tilde{I}_1^2}{2}\right)\exp(I_2 - \tilde{I}_2).$$

Let $g(y) = \exp(-(I_1^2 - \tilde{I}_1^2)/2)\exp(I_2 - \tilde{I}_2)$. Since $\tilde{I}_1/I_1 \to \tau_{12}\tau_{22}^{-1/2}(\alpha'\xi)^{-1}(\alpha'\alpha)^{1/2}$, $g(y)$ has a finite non-zero limit and is denoted by g_0. Applying L'Hospital's rule to $\exp(I_2 - \tilde{I}_2) = \exp(\xi'\xi/2)/\exp(\tau_{11}/2)$,

$$g(y) = \frac{\xi'\xi}{\tau_{11}}g(y) + \varepsilon(y)\exp(-(I_1^2 - \tilde{I}_1^2)/2), \tag{34}$$

where $\varepsilon(y)$ is a function of y and tends to 0 as $y \to \infty$. Suppose $\xi'\xi > \tau_{11}$. Then by (33), $\exp(-(I_1^2 - \tilde{I}_1^2)/2) \to 0$. We obtain $g_0 = (\xi'\xi/\tau_{11})g_0$, implying $\xi'\xi = \tau_{11}$, a contradiction to the assumption. Similarly we cannot have $\xi'\xi < \tau_{11}$. We thus establish $\xi'\xi = \tau_{11}$ in the case $\alpha'\xi < 0$, $\tau_{12} < 0$. Hence (30) reduces to

$$\Phi(I_1)/\Phi(\lambda/c_0) = \Phi(\tilde{I}_1)/\Phi(\tilde{\lambda}/\tilde{c}_0). \tag{35}$$

We also have

$$\zeta_1(I_1)I_1^{(1)} = \zeta_1(\tilde{I}_1)\tilde{I}_1^{(1)}. \tag{36}$$

If $\alpha'\alpha = 0$, then the left-hand side of (35) equals to 1 so that $\tilde{I}_1 = \tilde{\lambda}/\tilde{c}_0$. This implies that $\tau_{22} = 0$. The converse is also true. In this case the distribution of $\mathbf{z}'A\mathbf{z}$ does not depend on λ and the distribution of $NG_1(k_1/2, \tau, \tilde{\lambda}; \phi_n)$ does not depend on $\tilde{\lambda}$. The moment generating function of Q in (25) is that of the noncentral chi-squares distribution $\chi^2_{k_1}(\xi'\xi)$.

Suppose $\alpha'\alpha > 0$ and $\tau_{22} > 0$. If $\lambda = 0$ and $\alpha'\xi = 0$, then $I_1^{(1)} = 0$ so that $\tilde{I}_1^{(1)} = 0$ by (36), implying $\tilde{\lambda} = 0$ and $\tau_{12} = 0$. Conversely, $\tilde{\lambda} = 0$ and $\tau_{12} = 0$ imply $\lambda = 0$ and $\alpha'\xi = 0$. In this case $I_1 = \tilde{I}_1 = 0$. The distribution of $\mathbf{z}'A\mathbf{z}$ does not depend on $\alpha'\alpha$ and the distribution of $NG_1(k_1/2, \tau, \tilde{\lambda}; \phi_n)$ does not depend on τ_{22}. We have $Q \sim \chi^2_{k_1}(\xi'\xi)$.

Suppose now $\lambda \neq 0$ or $\alpha'\xi \neq 0$ (with $\alpha'\alpha > 0$). Then $I_1^{(1)}$ and $\tilde{I}_1^{(1)}$ are not identical to 0 and by (35) and (36),

$$\exp\left(-\frac{I_1^2 - \tilde{I}_1^2}{2}\right)\frac{I_1^{(1)}}{\tilde{I}_1^{(1)}} = \frac{\Phi(I_1)}{\Phi(\tilde{I}_1)} = \frac{\Phi(\lambda/c_0)}{\Phi(\tilde{\lambda}/\tilde{c}_0)}. \tag{37}$$

Again applying Laha's argument with

$$h(y) = \frac{(I_1^{(1)})^2}{(\tilde{I}_1^{(1)})^2} - \frac{(\Phi(\lambda/c_0))^2}{(\Phi(\tilde{\lambda}/\tilde{c}_0))^2}\exp(I_1^2 - \tilde{I}_1^2), \tag{38}$$

we obtain that $(I_1^{(1)})^2/(\tilde{I}_1^{(1)})^2$ and $I_1^2 - \tilde{I}_1^2$ are constants, see Driscoll and Gundberg [9, p. 68]. Hence $I_1/\tilde{I}_1 = \tilde{I}_1^{(1)}/I_1^{(1)}$. Expand $I_1^2 - \tilde{I}_1^2$ with numerator

$$\begin{aligned}&\{\tau_{22}(\alpha'\xi)^2 - \alpha'\alpha\tau_{12}^2\}y^3\\ &\quad+\{2\lambda\alpha'\xi\tau_{22} + (\alpha'\xi)^2(1+\tau_{22}) - 2\tilde{\lambda}\tau_{12}\alpha'\alpha - \tau_{12}^2(1+\alpha'\alpha)\}y^2\\ &\quad+\{\lambda^2\tau_{22} + 2\lambda\alpha'\xi(1+\tau_{22}) - \tilde{\lambda}^2\alpha'\alpha - 2\tilde{\lambda}\tau_{12}(1+\alpha'\alpha)\}y\\ &\quad+\{\lambda^2(1+\tau_{22}) - \tilde{\lambda}^2(1+\alpha'\alpha)\}\end{aligned} \tag{39}$$

and denominator

$$\{1 + \alpha'\alpha(1+y)\}\{1 + \tau_{22}(1+y)\}. \tag{40}$$

The coefficient of y^3 in (39) must be 0, implying

$$\tau_{22}(\alpha'\xi)^2 = \alpha'\alpha\tau_{12}^2. \tag{41}$$

Hence $\alpha'\xi = 0$ if and only if $\tau_{12} = 0$. In this case the coefficient of y^2 in (39) is 0 so that its all coefficients must be zeros since the coefficient of y^2 (40) is not equal to 0. Thus $\lambda^2\tau_{22} = \tilde{\lambda}^2\alpha'\alpha$ and $\lambda^2(1+\tau_{22}) = \tilde{\lambda}^2(1+\alpha'\alpha)$. It follows that $\lambda^2 = \tilde{\lambda}^2$. Let $y \to \infty$, then

$$\frac{(I_1^{(1)})^2}{(\tilde{I}_1^{(1)})^2} \to \begin{cases} \frac{(\alpha'\xi)^2\tau_{22}}{\alpha'\alpha\tau_{12}^2} = 1 & \text{if } \alpha'\xi \neq 0, \tau_{12} \neq 0, \\ \frac{\lambda^2\tau_{22}}{\alpha'\alpha\tilde{\lambda}^2} = 1 & \text{if } \alpha'\xi = \tau_{12} = 0, \lambda \neq 0, \tilde{\lambda} \neq 0. \end{cases} \tag{42}$$

By (37), $I_1^{(1)}$ and $\tilde{I}_1^{(1)}$ have the same sign. Hence $I_1/\tilde{I}_1 = 1$.

Expand I_1 at $y = 0$ as

$$\begin{aligned} I_1 = c_0^{-1}\{\lambda + (-\frac{a\lambda}{2} + \alpha'\xi)y + (\frac{3a^2\lambda}{8} - \frac{a\alpha'\xi}{2})y^2 \\ + \left(-\frac{5a^3\lambda}{16} + \frac{3a^2\alpha'\xi}{8}\right)y^3 + \cdots\}, \end{aligned} \tag{43}$$

where $a = \alpha'\alpha(1 + \alpha'\alpha)^{-1}$. Define $\tilde{a}$ similarly and expand $\tilde{I}_1$. Compare the coefficients of $I_1 = \tilde{I}_1$. If $\lambda = 0$, it is easy to see that $\tilde{\lambda} = 0$, $\alpha'\xi = \tau_{12}$ and $\alpha'\alpha = \tau_{22}$. If $\lambda \neq 0$, then $\tilde{\lambda} \neq 0$. Let $b = \alpha'\xi/\lambda$ and $\tilde{b} = \tau_{12}/\tilde{\lambda}$, we obtain

$$\lambda/c_0 = \tilde{\lambda}/c_0, \tag{44}$$

$$-\frac{a}{2} + b = -\frac{\tilde{a}}{2} + \tilde{b}, \tag{45}$$

$$\frac{3a^2}{8} - \frac{ab}{2} = \frac{3\tilde{a}^2}{8} - \frac{\tilde{a}\tilde{b}}{2}. \tag{46}$$

The solutions of (45) and (46) are $\tilde{a} = a$ and $\tilde{a} = 4b - 3a$, $\tilde{b} = b - (a - \tilde{a})/2$. Substituting $\tilde{a} = 4b - 3a$ into the equality of the coefficients of y^3 in I_1 and $\tilde{I}_1$, we obtain $(\tilde{a} - a)^3/32 = 0$. Thus $\tilde{a} = a$ is the unified expression of the solution, which implies $\alpha'\alpha = \tau_{22}$ and $\tilde{b} = b$. These equalities combined with (44) imply that $\lambda = \tilde{\lambda}$ and $\alpha'\xi = \tau_{12}$. □

Proof of Theorem 1. Without loss of generality we suppose $\Omega = I$.

Sufficiency: By assumption, there exists an orthogonal matrix Γ such that $\Gamma' A_i \Gamma$ diagonal with the $\sum_{j=1}^{i-1} k_j + 1$-th to $\sum_{j=i-1}^{i} k_j$-th diagonal elements being 1, others 0, ($\sum_{j=1}^{0} k_j = 0$). Let $\mathbf{y} = \Gamma'\mathbf{z}$. Then $\mathbf{y} \sim S_k(\Gamma'\xi, I, \lambda, \Gamma'\alpha; \phi_n)$. Partition $\Gamma = (\Gamma_1, \ldots, \Gamma_h)$, where Γ_i is $k \times k_i$. Partition $\mathbf{y} = (\mathbf{y}_1', \ldots, \mathbf{y}_h')'$. By Definition 1 $(Q_1, \ldots, Q_h) = (\mathbf{y}_1'\mathbf{y}_1, \ldots, \mathbf{y}_h'\mathbf{y}_h) \sim NG_h(k_1/2, \ldots, k_h/2; \tau, \lambda; \phi_n)$, where $\tau_{i,11} = (\Gamma_i'\xi)'\Gamma_i'\xi = \xi' A_i \xi$, $\tau_{i,12} = (\Gamma_i'\xi)'\Gamma_i'\alpha = \xi' A_i \alpha$, $\tau_{i,22} = (\Gamma_i'\alpha)'\Gamma_i\alpha = \alpha' A_i \alpha$.

Necessity: By the basic property of the *NG* distribution, marginally $\mathbf{z}' A_i \mathbf{z}$ is $NG_1(k_i/2; \tilde{\tau}, c\tilde{\lambda}; \phi_{k_i+1})$ for suitable parameters $\tilde{\tau}$ and c. Also by the reproductive property $\sum_{i=1}^{h} \mathbf{z}' A_i \mathbf{z}$ is $NG_1(\sum_{i=1}^{h} k_i/2; \sigma, \tilde{\lambda}; \phi_n)$ for suitable parameter σ. Hence by Lemma 1, $A_i^2 = A_i$, $\text{rank}(A_i) = k_i$, $i = 1, \ldots, h$, and $(\sum_{i=1}^{h} A_i)^2 = \sum_{i=1}^{h} A_i$. These conditions also imply that $A_i A_j = 0$ for $i \neq j$, $i, j = 1, \ldots, h$, see Anderson and Styan [3]. The parameters are determined using the marginal distributions by Lemma 1. □

Proof of Theorem 2. Note under the assumption of $\sum_{i=1}^{h} A_i = I$, $\text{rank}(A_i) = k_i$, the three conditions that $A_i^2 = A_i$ $(i = 1, \ldots, h)$, $\sum_{i=1}^{h} k_i = k$ and $A_i A_j = 0$ $(i \neq j$, $i = 1, \ldots, h)$ are equivalent, see Anderson and Styan [3]. Hence Theorem 1 leads to the sufficient and necessary conditions of Theorem 2. By (10), the moment generating function of $(Q_1, \ldots, Q_h)$ is

$$M(t) = \Phi\left(\frac{\lambda + \sum_{i=1}^{h} \alpha' A_i \xi 2t_i (1-2t_i)^{-1}}{(1+\sum_{i=1}^{h} \alpha' A_i \alpha t_i (1-2t_i)^{-1})^{\frac{1}{2}}}\right) \times \exp\left(\sum_{i=1}^{h} \xi' A_i \xi t_i (1-2t_i)^{-1}\right) \prod_{i=1}^{h} (1-2t_i)^{-\frac{k_1}{2}} / \Phi(\lambda/c_0). \tag{47}$$

If $A_i\alpha \neq 0$ for at most one i (say $i = 1$), then the first factor $\Phi(\cdot)$ in $M(t)$ does not depend on $t_2, \ldots, t_h$, and $M(t)$ is a product of the moment generating functions of Q_i with the specified distribution. □

7. Discussion

In this paper we introduce a version of the generalized Dirichlet distribution as the distribution of the noncentral quadratic form of the skew elliptical vector. This generalizes distributions of the quadratic forms studied in the literature. General form of the distribution and probability density function are obtained. Simpler form of the probability density and moment generator function are also obtained when the underlying distribution is skew normal. Our main result is the extension of Cochran's Theorem, which has been studied extensively in the literature of multivariate analysis. Two versions of Cochran's Theorem are given on the basis of Lemma 1. The sufficiency of these theorems provides fundamental conditions for applying the distribution in practice. The formulas of moments are used to obtain an estimate of the multivariate kurtosis for a real data set. We also provide an example of least-squares estimate where this distribution occurs and compare its property with that under normality. Note there are cases where the sufficient conditions of these theorems do not hold. For example, the matrix A to form the quadratic form $\mathbf{z}'A\mathbf{z}$ for estimating the autocovariance function in a time series model in Genton et al. [13, Eq. (10)] does not satisfy $A^2 = A$. Further research on more delicate distribution theory is desired. The proof of necessity of Lemma 1 is much more difficult than the non-skew case because of the skewness parameters presenting in the additional term $\Phi(I_1)/\Phi(\lambda/c_0)$ in the moment generating function (25). In the literature of the non-skew case the argument of Laha [15], see also Driscoll and Gundberg [9, p. 67], plays an important role in providing a correct proof of Craig's Theorem. Applying Laha's argument in (29) ensures identification of A and the noncentrality parameter τ_{11} in non-skew case. However, in the skew case we can only identify A at this stage to obtain (30) and more efforts are made to identify τ_{11} and other parameters. In addition to the property of the moment generating function, the property of its first order derivative is used. We use asymptotical argument to identify the sign of the parameter τ_{12} and then identify τ_{11}. The most difficult case is $\tau_{12} < 0$. With (35) and (36) obtained, we apply Laha's argument for the second time in (38) and some other arguments to identify the rest parameters λ, τ_{12} and τ_{22}. Alternative proof of the necessity using the equalities of the moments in the skew case seems not feasible, since the equations involve several parameters in a complicated way. It is pointed out in Section 4 that the sufficiency of the Cochran's Theorem hold for all density generator *f*. We conjecture that the necessity of the Cochran's Theorem in the skew case hold for a broad class of the density generator function f.

Acknowledgments

The author is thankful to two reviewers for their valuable comments and suggestions.

References

[1] M. Abramowitz, I.A. Stegun, Handbook of Mathematical Functions, Dover publications, New York, 1972.
[2] T.W. Anderson, K.T. Fang, Cochran's Theorem for elliptically contoured distributions, Sankhyā A 49 (1987) 305–315.
[3] T.W. Anderson, G.P.H. Styan, Cochran's Theorem, rank additivity and tripotent matrices, in: G. Kallianpur, P.R. Krishnaiah, J.K. Ghosh (Eds.), Statistics and Probability: Essays in Honor of C.R. Rao, North-Holland Publishing Company, Amsterdam, 1982, pp. 1–23.
[4] B.C. Arnold, R.J. Beaver, Hidden truncation models, Sankhyā 62 (2000) 23–25.
[5] A. Azzalini, A. Dalla Valle, The multivariate skew-normal distribution, Biometrika 83 (1996) 715–726.
[6] A. Azzalini, A. Capitanio, Statistical applications of the multivariate skew normal distribution, J. Roy. Statist. Soc. B 61 (1999) 579–602.
[7] M.D. Branco, D.K. Dey, A general class of multivariate skew-elliptical distributions, J. Multivariate Anal. 79 (2001) 99–113.
[8] R.D. Cook, S. Weisberg, An Introduction to Regression Graphics, Wiley, New York, 1994.
[9] M.F. Driscoll, W.R. Gundberg, A history of the development of Craig's Theorem, Amer. Statist. 40 (1986) 65–70.
[10] J. Fan, Distributions of quadratic forms and non-central Cochran's Theorem, Acta Math. Sinica 2 (1986) 185–198.
[11] B.Q. Fang, The skew elliptical distributions and their quadratic forms, J. Multivariate Anal. 87 (2003) 298–314.
[12] K.T. Fang, S. Kotz, K.W. Ng, Symmetric Multivariate and Related Distributions, Chapman & Hall, London, 1990.
[13] M.G. Genton, L. He, X. Liu, Moments of Skew-normal random vectors and their quadratic forms, Statist. Probab. Lett. 51 (2001) 319–325.
[14] R.C. Gupta, N. Brown, Reliability studies of the skew-normal distribution and its application to a strength-stress model, Commun. Statist. Theory Methods 30 (2001) 2427–2445.
[15] R.G. Laha, On the stochastic independence of two second-degree polynomial statistics in normally distributed variates, Ann. Math. Statist. 27 (1956) 790–796.
[16] N. Loperfido, Quadratic forms of skew-norm random vectors, Statist. Probab. Lett. 54 (2001) 381–387.
[17] K.V. Mardia, Measures of multivariate skewness and kurtosis with applications, Biometrika 57 (1970) 519–530.
[18] R.J. Muirhead, Aspects of Multivariate Statistical Theory, Wiley, New York, 1982.
[19] C.R. Rao, Linear Statistical Inference, 2nd ed., Wiley, New York, 1973.